


Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Merch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Sat May 13 10:39:20 2000; Maspar time 4.08 Seconds
 Tabular output not generated. 273.006 Million cell updates/sec

Title: >US-09-331-631-7
 Description: (34-80) from US09331631.pep (2 of 3)
 Perfect Score: 361
 Sequence: 1 YERDPROQYEQCQRCSEATEEREOQCRCEREXKEDQRODEE 47

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a:geneseq35
 1:geneseqp

Statistics: Mean 22.631; Variance 99.788; scale 0.227

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description	Prod. No.
1	361	100.0	525	1	Theobroma cacao antimi	1.51e-24
2	361	100.0	566	1	Sequence encoded by 67	1.51e-24
3	216	59.8	625	1	Macadamia integrifolia	5.45e-11
4	214	59.3	666	1	Macadamia integrifolia	1.02e-10
5	211	58.4	666	1	Macadamia integrifolia	4.93e-08
6	181	50.1	590	1	Gossypium hirsutum ant	1.72e-01
7	105	29.1	28	1	Stenocarpus sinuatus a	6.40e-01
8	98	27.1	1311	1	Precis coenia (butterf	6.40e-01
9	98	27.1	1311	1	Precis coenia patched	7.72e-01
10	97	26.9	162	1	Sequence of a new cyto	9.30e-01
11	96	26.6	186	1	Trypanosoma cruzi anti	1.95e+00
12	94	26.0	181	1	AdunAP45.	1.95e+00
13	92	25.5	76	1	AdunAP45.	1.95e+00
14	92	25.5	180	1	AdunAP45.	1.95e+00
15	91	25.2	181	1	AdunAP45.	2.34e+00
16	89	24.7	86	1	GST-HD fusion protein	3.38e+00
17	88	24.7	86	1	GST-HD fusion protein	3.38e+00
18	89	24.7	94	1	GST-HD fusion protein	3.38e+00
19	89	24.7	94	1	GST-HD fusion protein	3.38e+00
20	87	24.1	263	1	Moloney murine leukemia	4.87e+00
21	87	24.1	537	1	Osteoinductive retrovi	4.87e+00
22	87	24.1	562	1	Leucocytozoan protozoa	4.87e+00
23	87	24.1	626	1	Peanut allergen Ara hi	4.87e+00

ID	Score	Query Match	Length DB	ID	Description	Prod. No.
24	85	23.5	190	1	R91710	7.00e+00
25	84	23.3	409	1	W90342	8.38e+00
26	84	23.3	489	1	W90341	8.38e+00
27	84	23.3	637	1	W62837	8.38e+00
28	83	23.0	593	1	W62835	1.00e+01
29	83	23.0	971	1	W48896	1.00e+01
30	83	23.0	1048	1	W27277	1.00e+01
31	83	23.0	1048	1	W27277	1.00e+01
32	82	22.7	76	1	R90555	1.20e+01
33	82	22.7	76	1	R90555	1.20e+01
34	82	22.7	83	1	R90546	1.20e+01
35	82	22.7	567	1	W88788	1.20e+01
36	81	22.4	614	1	W62834	1.44e+01
37	81	22.4	614	1	W22149	1.44e+01
38	80	22.2	303	1	W60054	1.72e+01
39	80	22.2	371	1	W73369	1.72e+01
40	78	21.6	432	1	W93954	2.45e+01
41	78	21.6	1235	1	W65158	2.45e+01
42	77	21.3	217	1	W46424	2.93e+01
43	77	21.3	705	1	R66597	2.93e+01
44	77	21.3	711	1	W82789	2.93e+01
45	77	21.3	711	1	W07692	2.93e+01

ALIGNMENTS

RESULT 1
 ID W62831 standard; Protein; 525 AA.
 AC W62831.
 DE 27-OCT-1998 (first entry)
 DT Theobroma cacao antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Theobroma cacao.
 PN W09827805-A1.
 PD 02-JUN-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1: Page 47-49; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 CC be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 525 AA:
 Query Match 100.0%; Score 361; Length 525;
 Best Local Similarity 100.0%; Prod. No. 1.51e-24;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 34 YERDPROQYEQCQRCSEATEEREOQCRCEREXKEDQRODEE 80
 QY 34 YERDPROQYEQCQRCSEATEEREOQCRCEREXKEDQRODEE 80
 ID R20181 standard; Protein; 566 AA.
 AC R20181.
 DE 16-APR-1992 (first entry)
 DT Sequence encoded by 67 kb T. cacao protein cDNA.
 KW Cocoa; flavour; vicillin; seed storage protein.
 OS Theobroma cacao.
 PN W0919801-A.
 PD 26-DEC-1991.
 PF 07-JUN-1991; G00914.
 PR 11-JUN-1990; GB-013016.
 PA (MRSC) MARS UK LTD.
 PI Spencer ME, Hodge R, Deakin EA, Ashton S;
 DR WPI: 92-024418/03.
 DR N-PSDB: Q20377.
 PT Recombinant cocoa proteins - are responsible for flavour in cocoa

PT beans and produced in large quantities using yeast and bacterial expression vectors.
 PS Claim 4: Fig 2: 59pp; English.
 CC The inventors claim a 67 kd and 31 kd T. cacao protein, and fragments, and encoding DNAs. The 47 kd and 31 kd proteins are derived from the 67 kd precursor. T. cacao protein cDNA was detected in a cDNA library prepared from immature cocoa beans RNA using a probe based on the AA sequence of a CNR peptide common to the 47 kd and 31 kd polypeptides. Homology searches revealed close CC homologies between the 67 kd polypeptide and the vicilins, which are CC seed storage proteins.
 SQ Sequence 566 AA;

Query Match 100.0%; Score 361; DB 1; Length 566;
 Best Local Similarity 100.0%; Pred. No. 1.51e-24;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 YERDPRQOYECCORRCSEATEEREOECORCERERKEQORQEE 80
 QY 34 YERDPRQOYECCORRCSEATEEREOECORCERERKEQORQEE 80

RESULT 3
 ID W62830 standard; Protein; 625 AA.
 AC W62830.
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1..28 "signal peptide"
 FT Protein 29..666 /note="mature protein"
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB: V42316.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 43-45; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian CC animals.
 SQ Sequence 625 AA;

Query Match 59.8%; Score 216; DB 1; Length 625;
 Best Local Similarity 59.1%; Pred. No. 3.59e-11;
 Matches 26; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

Db 78 QRDPOQOYECCORRCSEATEEREOECORCERERKEQORQ 121
 QY 35 ERDPRQOYECCORRCSEATEEREOECORCERERKEQORQ 77

RESULT 4
 ID W62829 standard; Protein; 666 AA.
 AC W62829.
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1..28
 FT Protein 29..666 /note="signal peptide"
 FT PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB: V42316.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 43-45; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian CC animals.
 SQ Sequence 666 AA;

PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB: V42316.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 39-41; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian CC animals.
 SQ Sequence 666 AA;

Query Match 59.3%; Score 214; DB 1; Length 666;
 Best Local Similarity 59.1%; Pred. No. 5.45e-11;
 Matches 26; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

Db 119 QRDPOQOYECCORRCSEATEEREOECORCERERKEQORQ 162
 QY 35 ERDPRQOYECCORRCSEATEEREOECORCERERKEQORQ 77

RESULT 5
 ID W62828 standard; Protein; 666 AA.
 AC W62828.
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1..28 "signal peptide"
 FT Protein 29..666 /note="mature protein"
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB: V42316.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 34-36; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian CC animals.
 SQ Sequence 666 AA;

Query Match 58.4%; Score 211; DB 1; Length 666;
 Best Local Similarity 56.8%; Pred. No. 1.02e-10;
 Matches 25; Conservative 9; Mismatches 9; Indels 1; Gaps 1;

Db 119 QRDPOQOYECCORRCSEATEEREOECORCERERKEQORQ 162
 QY 35 ERDPRQOYECCORRCSEATEEREOECORCERERKEQORQ 77

RESULT 6
 ID W62832 standard; Protein; 590 AA.
 AC W62832.
 DT 27-OCT-1998 (first entry)
 DE Gossypium hirsutum antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Gossypium hirsutum.
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB: V42316.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 34-36; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian CC animals.
 SQ Sequence 666 AA;

OY 42 YECCORCESEATEEREEOCEOR-C 66

RESULT 15

ID R91712 standard; Protein: 181 AA.
AC R91712.
DT 17-NOV-1996 (first entry)
DE AcANAP47.
KW AcANAP; HONAP; NamAP; AcenAP; ADUNAP; anticoagulant;
KW nematode-extracted anticoagulant protein; serine protease;
KW nematode; thrombosis; parasitic worm.
OS Ancylostoma caninum.
PN WO9612021-A2.
PD 25-APR-1996.
PF 17-OCT-1995; U13231.
PR 18-OCT-1994; US-326110.
PR 05-JUN-1995; US-486389.
PR 05-JUN-1995; US-461965.
PR 05-JUN-1995; US-486397.
PR 05-JUN-1995; US-465380.
PA (CORV-) CORVAS INT INC.
PI Bergum PM, Ganssems VCI, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JH, Moyle W, Stanssens PEH;
PI Vlaux GP.
DR MPI; 96-222007/22.
DR N-PSDB; T12958.
PT Proteins with anticoagulant and/or serine protease inhibitory
PT activity - isolated from nematodes and useful to inhibit blood
PT coagulation
PS Claim 221; Fig 13G; 243pp; English.
CC Proteins with anticoagulant and/or serine protease inhibitory
CC activity, isolated from nematodes, are useful to inhibit blood
CC coagulation. The proteins can be added to blood collection tubes
CC defining the collection of mammalian plasma. They are also useful
CC to prevent or inhibit thrombosis, and may be given alone or in
CC combination with other therapeutic or in vivo diagnostic agents.
CC The proteins can serve as immunogens to raise antibodies for use in
CC the diagnosis and identification of NAP concn. levels in biological
CC fluids, e.g. to detect mammalian infection with a parasitic worm.
CC They can also be used as immunogens in prophylactic and therapeutic
CC vaccines against parasitic worm infection. The proteins may
CC double the clotting time of human plasma in prothrombin time assays
CC when present at 10-50 nMol, and double the clotting time of human
CC plasma in activated partial thrombin time assays when present
CC at 10-100 nMol.
CC The anticoagulant proteins are pref. derived from
CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator
CC americanus or Heligmosomoides polygyrus.
CC The proteins pref. have 2 NAP domains and specifically inhibit
CC the catalytic activity of the factor VIIa/TF complex in the
CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
CC do not specifically inhibit the activation of factor VIIa in the
CC absence of TF and do not specifically inhibit prothrombinase.
SO Sequence 181 AA:

Query Match 25.2%; Score 91; DB 1; Length 181;

Best Local Similarity 46.2%; Pred. No. 2.34e+00;

Matches 12; Conservative 5; Mismatches 7; Indels 2; Gaps 2;

DB 118 YKCRKC-SELSKNEPACLSRAC 142

OY 42 YECCORCESEATEEREEOCEOR-C 66

Search completed: Sat May 13 10:39:28 2000
Job time : 8 secs.